

Abstract of KR2002-12645 :

PURPOSE: A web server operation system and method is provided to offer various information provision service e.g. voice, static picture or moving picture, according to a users request or a traffic priority by applying a traffic management technique in order to guarantee a quality of service and minimize a load in the web server.

CONSTITUTION: The system comprises a traffic classifier(300), a load measurer(305), a token generator(310), and a service scheduler(315). The traffic classifier(300) classifies the traffics generated when a user requests a current service to a web server. The load measurer(305) takes a real time measurement of a load amount of the web server. The token generator(310) generates a token corresponding to the classification of the traffic the user currently requests. The service scheduler(315) puts a priority on the services by the order of a guaranteed service, a controlled load service and a best-effort service, and processes the corresponding service with a higher priority.

(19) 대한민국특허청 (KR)
(12) 공개특허공보(A)

(51) Int. Cl.⁷
G06F 17/00

(11) 공개번호 특2002-0012645
(43) 공개일자 2002년02월20일

(21) 출원번호 10-2000-0045732
(22) 출원일자 2000년09월08일
(71) 출원인 주식회사 이지컴유니케이션즈 지계만
서울 강남구 논현1동 206-6
(72) 발명자 유민터
경기도성남시분당구서현동헌대마파드413동 402호
(74) 대리인 최미옥, 이경란

실용신안 : 없음

(54) 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템 및 방법

요약

본 발명은 효율적인 웹 서버 운용을 통하여 효과적인 웹 서비스를 제공하는 통합 웹 서비스를 위한 인터넷 웹 서버 운용 방법을 관한 것이다. 본 발명에 따른 통합 웹 서비스를 위한 인터넷 웹 서버 운용 방법은 웹 서비스를 위한 웹 서버 운용 방법에 있어서, 사용자로부터 웹 서비스 요청 정보를 수신하고, 수신된 웹 서비스의 종류를 판별하고, 판별 결과, 미리 설정된 서비스의 우선 순위에 따라 해당 웹 서비스를 수행한다. 따라서, 본 발명에 의하면 인터넷 상에서 웹 서비스를 제공하고 있는 웹 서버의 운용을 실시간 트래픽 감시 정보를 기초로 하여 트래픽 관리 기능을 적용함으로써 소프트웨어적 혹은 알고리즘적으로 해결할 수 있고, 지속적으로 증가하는 웹 서비스에 대한 차별화된 서비스를 요구하는 사용자 요구 사항에 부응하여 사용자에게 향상된 품질의 서비스를 제공할 수 있다.

도면

도 3

색인어

트래픽, 부하, 스케줄러, 인터넷, 웹, 네트워크, 토론

참고문헌

도면의 간단한 설명

- 도 1은 인터넷 상에서 운영되는 웹 서버의 사용자 접근에 따른 상태 변화를 나타낸 도면이다.
도 2는 본 발명의 바람직한 다른 실시예에 따른 웹 서버의 상태에 상응하여 제공되는 서비스들 표시한 도면이다.
도 3은 본 발명의 바람직한 웹 실시예에 따른 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템의 구성도.
도 4는 본 발명의 바람직한 다른 실시예에 따른 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템의 구성도.
도 5a는 본 발명의 바람직한 일 실시예에 따른 사용자로부터 서비스 요구시에 서비스 우선 순위를 적용한 웹 서버 운용 시스템의 데이터 처리를 간략하게 나타낸 순서도.
도 5b는 본 발명의 바람직한 일 실시예에 따른 사용자로부터 서비스 요구시에 서비스 우선 순위를 적용한 웹 서버 운용 시스템의 데이터 처리를 상세하게 나타낸 도면.
도 6a는 본 발명의 바람직한 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 서비스 요구 순서대로 데이터 처리를 간략하게 나타낸 도면.
도 6b는 본 발명의 바람직한 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 서비스 요구 순서대로 데이터 처리를 상세하게 나타낸 도면.
도 7a는 본 발명의 바람직한 또 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 요구 서비스에 우

선 순위를 적용한 데이터 처리를 강력하게 나타낸 도면.

도 7a와 도 7b는 본 발명의 바람직한 또 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 요구 시 데이터 우선 순위를 적용한 데이터 처리를 상세하게 나타낸 도면.

<도면의 주요 부분에 대한 부호의 설명>

100 : 트래픽 관리부 105 : 웹 서버의 부하 감시 장치부

110 : 트래픽 부류별 토큰 발생부 115 : 서비스 스케줄러부

발명의 상세한 설명

발명의 목적

발명이 속하는 기술분야 및 그 분야의 종래기술

본 발명은 멀티미디어 서비스를 제공하는 웹 서버에서의 서비스 제공에 따른 웹 서버를 효과적으로 운용하는 통합 웹 서비스를 위한 인터넷 상에서의 웹 서버 운용 시스템 및 방법에 관한 것이다.

종래에는 인터넷 상을 통한 클라이언트의 서버에의 접근은 현재 무수히 많이 이루어지고 있으나, 네트워크 측면에서의 트래픽 관리 기법이 현재 단일 시스템의 고성능 웹 서비스를 제공하기 위한 방법으로 사용되고 있지 않았다.

여기서, 종전의 트래픽 관리 기법이란 토큰 링 방식을 의미하는 것으로서, 토큰 링(token ring)은 IBM 케이블링 시스템(cabling system)에 사용된, IBM에 의하여 발명한 링(LAN:국간기-통신망)기술이다. 링 전송에 대한 액세스 방식의 하나로, IEEE(Institute of Electrical and Electronics Engineers)802.5에 방법을 둔 링이나, FDDI(Fiber Distributed Data Interface)가 이 방식을 채택하고 있다.

토큰 링 방식에서는 워크스테이션 등의 노드(node)를 전송로에서 고리 모양으로 접속한다. 링 위는 토큰이라고 하는 특수한 데이터가 한 개 순회한다. 그리고 토큰을 수취한 노드가 데이터를 송신할 수 있으며, 데이터를 송신한 노드는 자신이 송신한 데이터가 링을 한 바퀴 돌고 돌아오면, 그것을 링으로부터 제거한다. 송신할 데이터가 없는 노드는 토큰을 다음 노드에 보낸다.

이러한 토큰 링 방식은 네트워크에서의 각 노드의 접속 방식 중 하나이다. 토큰 링 방식에서는 네트워크를 토큰(loop)로 구성하고, 토큰이라고 하는 액세스권(access rights)을 나타내는 신호가 네트워크를 돌고, 송신을 하는 노드는 토큰을 얻어서 제어권을 획득하면 메시지 패킷(packet)을 전달한다.

한편, 웹 서버의 고속화, 대용량화, 그리고 고성능화가 빠른 속도로 진행되어 현재까지 무한한 웹 서비스가 진행되고 있다.

그러나, 네트워크 측면에서의 전송 링크의 고속화가 네트워크 내부의 트래픽 과부하에 대한 근본적인 해결책이 되지 못하는 문제점이 있다.

그리고 웹 서비스가 하드웨어적 구성 요소의 고성능화뿐만 제공되고 있기 때문에 멀티미디어 인터넷 트래픽 지원을 위한 근본적인 해결책이 되지 못하는 문제점이 있다.

발명이 이루고자 하는 기술적 과제

본 발명은 상기의 문제점을 해결하기 위하여, 제안된 것으로서, 네트워크 상에서 이루어졌던 트래픽 관리 기법을 인터넷 서비스를 위한 웹 서버 운영에 적용함으로써 효율적인 인터넷 서비스를 제공하고자 하는 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템 및 방법이다.

따라서, 본 발명의 목적은 웹 서버에서의 실시간 트래픽 감시 정보에 기초한 웹 서버 운용 방법을 적용함으로써 효율적인 인터넷 서비스를 제공할 수 있는 통합 웹 서비스를 위한 웹 서버 운용 시스템을 제공하는 것이다.

본 발명의 다른 목적은 웹 서버 내에서 제공하는 음성 정보, 정치 영상 정보, 동영상 정보 등의 다양한 서비스를 사용자 요구 및 트래픽 우선 순위에 따라 제공함으로써 사용자에게 서비스 품질을 보장함과 동시에 웹 서버 내의 부하를 최소화 할 수 있는 통합 웹 서비스를 위한 웹 서버 운용 시스템을 제공하는 것이다.

발명의 구성 및 작용

상기 목적들을 달성하기 위하여, 본 발명의 일 측면에 따르면, 웹 서비스를 위한 웹 서버 운용 방법이 있어서, 사용자로부터 웹 서비스 요청 정보를 수신하고, 수신된 웹 서비스의 종류를 판별하여, 판별 결과, 미리 설정된 서비스의 우선 순위에 따라 해당 서비스를 수행하는 통합 웹 서비스를 위한 웹 서버 운용 방법, 상기 방법에 상응한 장치 및 시스템이 제공된다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 일 실시예에 따르면, 상기 웹 서비스는 보강형 서비스, 부하 제어형 서비스 그리고 인터넷 서비스로 구분될 수 있다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 다른 실시예에 따르면, 상기 서비스는 요청된 각 서비스마다 트론을 부여하고, 트론이 부여된 각 서비스에 대해 상기 설정된 우선 순위에 따라 수행될 수 있다.

상기 목적들을 달성하기 위하여, 본 발명의 다른 측면에 따르면, 웹 서비스를 위한 웹 서버 운용 방법이 있어서, 웹 서비스의 종류에 따라 우선 순위를 설정하고, 사용자로부터 웹 서비스 요청이 있는 경우 해당 웹 서비스를 수행하며, 웹 서비스 수행 중에 다른 종류의 웹 서비스 정보 요청이 있는 경우, 요청된 웹 서비스의 종류를 판단하여 진행 중인 웹 서비스와 현재 요청된 웹 서비스의 우선 순위를 비교 판단하고, 우선 순위가 높은 웹 서비스를 수행하는 동안 웹 서비스를 위한 웹 서버 운용 방법, 상기 방법에 상응한 장치 및 시스템이 제공된다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 일 실시예에 따르면, 상기 웹 서비스는 보강형 서비스, 부하 제어형 서비스 그리고 기존 인터넷 서비스로 구분될 수 있다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 다른 실시예에 따르면, 상기 서비스는 요청된 각 서비스마다 트론을 부여하고, 트론이 부여된 각 서비스에 대해 상기 설정된 우선 순위에 따라 수행될 수 있다.

상기 목적들을 달성하기 위하여, 본 발명의 또 다른 측면에 따르면, 복수의 서비스에 대해 각각 우선 순위를 설정한 후에, 사용자로부터 서비스 요청 정보를 수신하고, 수신한 서비스 요청 정보에 상응하여 트론을 부여하고, 부여된 트론에 해당하는 서비스를 제공하고, 상기 서비스의 제공 중에 상기 우선 순위의 트론이 존재하는 서비스의 요청이 있는지를 판단하며, 판단 결과, 제공 중인 서비스의 트론보다 현재 요청된 서비스의 트론이 우선 순위가 높은 경우, 우선 순위가 높은 해당 서비스를 우선적으로 사용자에게 제공하고, 이후에 제공 중인 서비스를 계속 수행하고, 상기 서비스를 제공한 후, 해당 서비스 정보의 트론을 제거하는 통합 웹 서비스를 위한 웹 서버 운용 방법, 상기 방법에 상응한 장치 및 시스템이 제공된다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 일 실시예에 따르면, 상기 결과 결과, 제공 중인 서비스의 트론이 현재 요청된 서비스의 트론보다 상기 우선 순위의 트론인 경우에는 제공 중인 서비스는 계속 수행될 수 있다.

통합 웹 서비스를 위한 웹 서버 운용 방법이 제공된 바람직한 다른 실시예에 따르면, 상기 서비스는 보강형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스로 구분될 수 있다.

이하, 첨부한 도면들을 참조하여 본 발명의 바람직한 실시예를 상세히 설명하기로 한다.

도 1은 인터넷 상에서 운영되는 웹 서버의 사용자 접근에 따른 상태 변화를 나타낸 도면이다.

도 1을 참조하면, 단계 100에서 웹 서버는 정상 상태에서 동작 시에 외부 및 내부의 클라이언트는 상용이 발생하지 않은 정상 상태를 계속 유지하게 된다. 또한 웹 서버는 정상 상태에서 웹 서버 사용자 수의 변동을 적을 경우 기존 인터넷 트래픽, 요청 장치 등을 이용하여 내부적으로 트래픽 관리 및 패킷 스케줄링을 수행함으로써 최적의 상태(단계 105)로 전이한다.

그러나 단계 100에서 웹 서버를 이용하는 사용자 수가 급증하는 경우라면, 웹 서버는 사용자 트래픽이 서버 하드웨어 용량을 초과하는 가상의 혼잡 상태(110)로 전이한다.

그리고 단계 105에서 웹 서버는 최적 상태에서 웹 서버를 이용하는 사용자 트래픽의 변동이 크지 않고 더 이상의 웹 서버에의 사용자 접근이 불가하지 않는 한 최적 상태를 계속 유지하며, 이후 웹 서버를 이용하는 사용자 수가 급증하는 경우에도 웹 서버는 최적의 상태에서 정상 상태(단계 100)로 전이한다.

단계 110에서 가상의 혼잡 상태는 웹 서버가 이미 정상적인 웹 서비스를 제공하기 힘든 상태에 해당하며, 이 경우 지속적으로 웹 서버에의 사용자 접근이 증가하게 되면, 웹 서버는 가상의 혼잡 상태에서 정상 상태(단계 100)로의 전이가 되지 않고 계속하여 가상의 혼잡 상태를 유지한다.

그러나 웹 서버는 정상 상태(단계 100)로의 전이를 위해 내부적으로 트래픽 관리 기능을 수행하거나 트래픽 플러스 플 패킷 스케줄링 방법 등을 적용한다.

또한, 단계 110에서 웹 서버는 가상의 혼잡 상태에서 계속되는 웹 서버 사용자 수의 증가가 발생하고 하더라도 구성 요소의 예측하지 못한 고장이 발생하면 실제의 혼잡 상태(단계 115)로 전이한다.

한편 웹 서버는 상기에서 설명한 정상 상태(단계 100), 최적의 상태(단계 105) 그리고 가상의 혼잡 상태(단계 110)로부터 예측하기 힘든 사용자 트래픽의 폭증이나 웹 서버 하드웨어 장치의 고장 등에 의하여 더 이상 사용자의 서비스 요구에 대하여 원하는 서비스 품질을 제공할 수 없는 상황에 도달했을 경우에 혼잡 상태(단계 115)에 이르게 된다.

도 1에 예시된 바와 같이 웹 서버는 정전하는 사용자 수의 증가 및 사용자 트래픽 특성 등에 따라 정상 상태(단계 100), 최적의 상태(단계 105), 가상의 혼잡 상태(단계 110) 그리고 혼잡 상태(단계 115)로 전이할 수 있는 네 가지 상태를 가지게 된다. 따라서 본 발명은 결과적으로 도 1에서 정의된 네 가지 상태 중에서 가능한 한 정상 상태 및 최적의 상태를 유지할 수 있도록 웹 서버를 운용하도록 하고자 하는 것이다.

도 2는 본 발명의 바람직한 다른 실시예에 따른 웹 서버의 상태에 상응하여 제공되는 서비스들을 표시한 도면이다.

도 2를 참조하면, 웹 서버는 사용자의 서비스 요구에 따른 상태를 도 1에 표시된 바와 같이 정상 상태(100 - 도 1 참조), 최적의 상태(105 - 도 1 참조), 가상의 혼잡 상태(110 - 도 1 참조) 그리고 혼잡

상대(115 - 도 1 참조)로 구분하고, 이와 같이 구분된 상태에 따라 제공 가능한 서비스의 종류를 미리 설정한다.

상기 내용을 상술하면, 웹 서버는 정상 상태(100 - 도 1 참조)와 최척 상태(105 - 도 1 참조)에서 사용자로부터 서비스 요구시에 도 2에 표시된 서비스 A(보장형 서비스), 서비스 B(부하 제어형 서비스) 그리고 서비스 C(가용 인터넷 서비스)가 제공된다.

여기서, 서비스 A는 보장형 서비스로서, 보장형 서비스는 "IETF(Internet Engineering Task Force) RFC(request for comments) 2212"에 정의되어 있는 서비스로서 이는 엄격한 QoS(Quality of Service)를 보장하는 서비스이다. 예를 들면, 정해진 데이터 레이트(Data rate)를 보장하고, 버퍼에 도달 레미(bounded delay), 노 큐잉 로스(no-queuing loss) 등의 특성을 가지고 있어, 하드 리얼 타임 애플리케이션(hard real-time applications) 또는 서킷 에뮬레이션(circuit emulation) 등에 적용되는 것으로 오디오 서비스나 비디오 서비스 등이 있을 수 있다.

서비스 B는 부하 제어형 서비스로서, 부하 제어형 서비스는 "RFC 2212"에 정의되어 있는 서비스이다. 예를 들면, 부하가 적을 때의 기존 인터넷 서비스와 근사한 서비스로서 대역폭, 지연 시간, 패킷 손실에 대한 절대적 보장은 없으나, 네트워크 내의 충분한 자원을 확보함으로써, 높은 전송 확률을 제공한다. 따라서 이 서비스는 어댑티브 리얼 타임 애플리케이션(Adaptive real-time applications) 등에 적용되는 것으로 화상기 예약 서비스 등과 같은 온라인 서비스 등이 있을 수 있다.

그리고 서비스 C는 기존 인터넷 서비스로서, 기존 인터넷 서비스는 일반적인 사용자 서비스 요구로서 파일 다운로드, 이메일 서비스 등이 있을 수 있다.

그리고 웹 서버는 가용의 혼잡 상태(110 - 도 1 참조)에서 사용자의 서비스 요구 시에 상기에서 설명한 서비스 A(보장형 서비스) 그리고 서비스 B(부하 제어형 서비스)를 제공한다.

혼잡 상태에서 웹 서버는 사용자의 서비스 요구 시에 상기에서 설명한 서비스 A(보장형 서비스)만을 제공한다.

이렇게 도 2와 같이 웹 서버의 상태에 따라 미리 제공 가능한 서비스가 설정됨에 따라 사용자의 요구에 따라 웹 서버는 제공 가능한 서비스만을 제공하게 되며, 이에 대한 설명은 도 8와 도 8b를 참조하여 상세히 설명하기로 한다.

도 3은 본 발명의 바람직한 일 실시예에 따른 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템의 구성도를 나타낸 도면이다.

도 3을 참조하면, 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템은 일반적인 서버 내의 시스템 구성(예를 들면, 수신펀, 입력부 등) 이외에 추가적으로 도 3에 예시된 바와 같이 웹서버에 현재 서비스의 요구하는 사용자의 트래픽을 분류하는 트래픽 분류부(300), 웹 서버의 부하량을 실시간으로 측정하는 웹서버의 부하 감지 장치부(305), 트래픽 분류 방법과 따라 현재 사용자가 요구하는 트래픽 분류에 해당하는 토큰(token)을 할당시키는 트래픽 분류부, 토큰 할당부(310) 그리고 3개의 독립 버퍼(가: 트래픽 부하 버퍼, 트: 트: 버퍼를 하나씩 관리함)를 가지고, 보장형(Guaranteed) 서비스, 부하 제어형(Controlled Load) 서비스, 기존 인터넷(Best-effort) 서비스의 순으로 우선 순위를 부여하여 상위 우선 순위의 토큰이 존재할 경우 항상 이들 우선순위로 처리하는 서버 스케줄러부(215)를 구비한다.

여기서 보장형 서비스는 "IETF(Internet Engineering Task Force) RFC(request for comments) 2212"에 정의되어 있는 서비스로서 이는 엄격한 QoS(Quality of Service)를 보장하는 서비스이다. 예를 들면, 정해진 데이터 레이트(Data rate)를 보장하고, 버퍼에 도달 레미(bounded delay), 노 큐잉 로스(no-queuing loss) 등의 특성을 가지고 있어, 하드 리얼 타임 애플리케이션(hard real-time applications) 또는 서킷 에뮬레이션(circuit emulation) 등에 적용되는 것으로 오디오 서비스나 비디오 서비스 등이 있을 수 있다.

부하 제어형 서비스는 "RFC 2212"에 정의되어 있는 서비스이다. 예를 들면, 부하가 적을 때의 기존 인터넷 서비스와 근사한 서비스로서 대역폭, 지연 시간, 패킷 손실에 대한 절대적 보장은 없으나, 네트워크 내의 충분한 자원을 확보함으로써, 높은 전송 확률을 제공한다. 따라서 이 서비스는 어댑티브 리얼 타임 애플리케이션(Adaptive real-time applications) 등에 적용되는 것으로 화상기 예약 서비스 등과 같은 온라인 서비스 등이 있을 수 있다.

그리고 기존 인터넷 서비스는 일반적인 사용자 서비스 요구로서 파일 다운로드, 이메일 서비스 등이 있을 수 있다.

상기와 같은 원리에 의하여 실시간 성이 필요하거나 높은 우선 순위를 가지는 서비스 요구에 대해서는 항상 우선순위가 고 서비스 종류에 보정되는 서비스를 제공할 수 있게 된다.

여기서 상기 실시간 성이 필요한 실시간 처리할 요구하는 트래픽을 의미하며, 상기와 같은 각 서비스의 구분은 IP(Internet Protocol) 프로토콜 헤더 내의 TOS(type of service) 필드를 이용한다.

이러한 상기와 같은 시스템에 운용하는 웹 서버 운용 방법을 간단히 설명하면, 서버는 복수의 서비스에 대해 각각 우선 순위를 설정하고, 사용자로부터 서비스 요청 정보를 수신하며, 수신한 서비스 요청 정보에 상응하여 토큰을 부여한다.

그리고 서버는 부여한 토큰에 해당하는 서비스를 제공하고, 상기 서비스의 제공 중에 상위 우선 순위의 토큰이 존재하는 서비스의 요청이 있는지를 판단한다. 이러한 판단이 있는 후, 제공 중인 서비스의 토큰보다 현재 요청된 서비스의 토큰이 우선 순위가 높은 경우, 우선 순위가 높은 해당 서비스를 우선적으로 사용자에게 제공한다. 이후에 제공 중인 서비스를 계속 수행하게 된다.

이 후, 서버는 상기 서비스를 제공한 후, 해당 서비스 정보의 토큰을 제거한다.

상술한 바와 같이 웹 서버 운용에 있어 대략적인 설명을 하였으나, 이 후, 도 5a를 통하여 더 상세히 설

명하기로 한다.

도 4는 본 발명의 바람직한 다른 실시예에 따른 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템의 구성도이다.

도 4를 참조하면, 통합 웹 서비스를 위한 인터넷 웹 서버 운용 시스템은 일반적인 서버 내의 시스템 구성(예를 들면, 수신부, 입력부 등) 이외에 추가적으로 도 4에 예시된 바와 같이 웹서버에 현재 서비스를 요구하는 사용자의 트래픽을 분류하는 트래픽 분류부(400), 그리고 서비스를 요구하는 사용자의 트래픽에 상응하여 토큰(token)을 부여하고, 해당 서비스의 완료 후에 부여된 토큰을 제거하는 도록, 발행부(405), 웹 서버의 부하량을 실시간으로 측정하는 웹서버의 부하 감시 장치부(410)(부하 감시 장치부에서 실시간으로 측정하는 부하량에 따라 웹 서버는 서버의 상태를 감작하게 된다.) 그리고

개의 토큰 비퍼 및 트래픽 부하 별로 토큰 비퍼를 하나씩 권리함)을 가지고, 보장형(Guaranteed) 서비스, 부하 제어형(Controlled load) 서비스, 기존 인터넷(Best-effort) 서비스의 우선 순위를 분기하여 상위 우선 순위의 토큰이 존재할 경우 항상 이를 우선적으로 처리하는 서비스 스케줄러부(415)를 구비한다.

상기 서비스의 종류에 대한 정의는 도 3를 참조하기로 한다.

상기와 같은 원리에 의하여 실시간 성이 필요하거나 높은 우선 순위를 가지는 서비스 요구에 대해서는 항상 안정적이고 서비스 품질이 보장되는 서비스를 제공할 수 있게 된다.

여기서 상기 실시간 성이란 실시간 처리를 요구하는 트래픽을 의미하며, 상기와 같은 각 서비스의 구분은 IP(Internet Protocol) 패킷 헤더 내의 TOS(type of service) 필드를 이용한다.

도 5a는 본 발명의 바람직한 일 실시예에 따른 사용자로부터 서비스 요구시에 서비스 우선 순위를 적용한 웹 서버 운용 시스템의 데이터 처리를 감작하게 나타낸 순서도이다.

도 5a를 참조하면, 단계 500에서 웹 서버는 서비스 종류를 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스를 분류하고, 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 순서로 우선 순위를 설정한다.

단계 505에서 사용자의 서비스 요구 시에 웹 서버는 이를 수신하고, 단계 510에서 웹 서버는 사용자로부터 요구된 해당 서비스를 수행한다. 예를 들면, 항공 예약 서비스와 같은 온라인 서비스 요구를 수신하고, 이에 대한 서비스 정보를 제공한다.

그리고 단계 515에서 웹 서버는 사용자로부터 요구된 해당 서비스를 수행하는 도중이라도 사용자로부터 새로운 서비스 요구(예를 들면, 파일 다운로드와 같은 우선 순위가 낮은 기존 인터넷 서비스)가 발생한다면, 새로운 사용자 요구 서비스가 현재 수행 중인 서비스보다 상위 우선 순위의 서비스인지 판단한다.

단계 515에서의 판단 결과, 새로운 서비스 요구가 상위 우선 순위의 서비스라면, 웹 서버는 현재 수행 중인 서비스를 중단하고 단계 510으로 다시 진행한다. 그러나, 상위 우선 순위의 서비스가 아니라면, 현재 수행 중인 서비스와 같은 종류의 서비스 또는 수행 중인 서비스보다 낮은 우선 순위의 서비스라면 단계 520으로 진행한다.

예를 들어 단계 515에 설명하면, 현재 수행 중인 서비스가 항공 예약과 같은 온라인 서비스이고, 사용자로부터 새로운 서비스 요구는 파일 다운로드와 같은 서비스 요구라고 가정한다.

상기의 가정에 따르면, 항공 예약과 같은 온라인 서비스는 부하 제어 서비스로서 파일 다운로드와 같은 기존 인터넷 서비스보다 상위 우선 순위의 서비스에 해당한다. 따라서 단계 515에서의 판단 결과 현재 수행 중인 항공 예약과 같은 온라인 서비스가 우선 순위가 더 높고, 따라서 단계 520으로 진행한다.

그러나, 상기의 예에서 사용자로부터 새로운 서비스 요구가 오디오 서비스와 같은 보장형 서비스라면, 이는 항공 예약과 같은 부하 제어 서비스보다 우선 순위가 높기에 현재 수행 중인 서비스를 중단하고, 단계 510으로 다시 진행한다.

단계 520에서 미리 설정된 우선 순위 즉, 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 순서로 서비스를 수행한다. 상기에서 설명한 예에 따르면, 새로운 서비스 요구가 상위 우선 순위의 서비스라면 새로운 요구 서비스를 수행하고, 설정된 우선 순위에 따라 현재 제공 중인 서비스보다 높은 우선 순위의 서비스 요구가 있을 경우에만 현재 제공 중인 서비스를 계속 수행한다.

도 5a는 설정된 우선 순위에 따른 서비스 제공의 경우를 간략히 설명한 것으로서 도 5a를 통하여 더 상세히 설명하기로 한다.

도 5a는 본 발명의 바람직한 일 실시예에 따른 사용자로부터 서비스 요구시에 서비스 우선 순위를 적용한 웹 서버 운용 시스템의 데이터 처리를 상세하게 나타낸 도면이다.

도 5b를 참조하면, 단계 525에서 웹 서버는 서비스 종류에 따라 미리 우선 순위를 설정한다. 즉, 모든 종류의 서비스를 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스로 분류하고, 사용자의 서비스 요구시 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 우선 순위에 따라 서비스를 제공한다.

여기서 보장형 서비스는 'IETF(Internet Engineering Task Force) RFC(request for comments) 2212'에 정의되어 있는 서비스로서 이는 엄격한 QoS(Quality of Service)를 보장하는 서비스이다. 예를 들면, 정해진 데이터 레이트(data rate)를 보장하고, 바운드된 딜레이(bounded delay), 노 큐잉 로스(no queueing loss) 등의 특성을 가지고 있어, 하드 리얼 타임 애플리케이션(hard real-time applications) 또는 시큐리티 애플리케이션(circuit emulation) 등에 적용되는 것으로 오디오 서비스나 비디오 서비스 등이 있을 수 있다.

부하 제어형 서비스는 'RFC 2212'에 정의되어 있는 서비스이다. 예를 들면, 부하가 적을 때의 기존 인터

넷 서비스와 근사한 서비스로서 매력적, 자연 시간, 패킷 손실에 대한 절대적 보장은 없으나, 네트워크 내의 충분한 자원을 확보함으로써, 높은 전송 속도를 제공한다. 따라서, 이 서비스는 이터티브 리얼 타임 애플리케이션(Adaptive real-time applications) 등에 적용되는 것으로 항공기 예약 서비스 등과 같은 온라인 서비스 중이 있을 수 있다.

그리고 기존 인터넷 서비스는 일반적인 사용자 서비스 요구로서 파일 다운로드, 이메일 서비스 등이 있을 수 있다.

단계 530에서 웹 서버는 사용자의 서비스 요구 정보를 수신하고, 수신한 서비스 요구 정보를 서버 내의 트래픽 분류부(300 - 도 3 참조)로 전달한다.

단계 535에서 트래픽 분류부(300 - 도 3 참조)는 수신한 사용자의 요구 정보에 따라서 트래픽을 분류한다. 이 때, 트래픽 분류부(300 - 도 3 참조)는 사용자의 서비스 요구에 따라 제공되는 서비스를 보장할 서비스, 부하 제어 서비스, 기존 인터넷 서비스 등으로 구분한다.

이러한 트래픽의 서버 내에서의 분류 방법은 패킷(Packet) 내 특정 필드의 해석으로 구현할 수 있다.

단계 540에서 트래픽 분류부(300 - 도 3 참조)에 의해 분류된 트래픽(사용자에 의한 요구 서비스)에 대하여 트래픽 부류별 토큰 발생부(310 - 도 3 참조)는 해당 서비스 요구에 상응한 트래픽에 토큰을 부여하고 (토큰 수를 증가시킴), 토큰을 서비스 스케줄러부(315 - 도 3 참조)로 전달한다.

단계 545에서 각 트래픽 부류(즉, 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스)에 해당하는 트래픽을 지정할 때로 토큰 배분을 하나씩 판단하고 있는 서비스 스케줄러부(315 - 도 3 참조)는 우선 순위로 보장형 서비스에 대한 토큰 배분을 검사하여 보장형 서비스에 해당하는 토큰이 존재하는지를 판단한다. 다시 말하면, 보장형 서비스에 대한 사용자 요구가 있었는지를 판단하는 것이다. 판단 결과, 보장형 서비스에 대한 사용자 요구가 있었다면, 단계 550으로 진행한다. 그러나 단계 545에서 서비스 스케줄러부(315 - 도 3 참조)는 보장형 서비스에 대한 토큰 배분을 검사하여 판단한 결과, 보장형 서비스의 토큰이 존재하지 않는다면 즉, 보장형 서비스에 대한 사용자 요구가 없다면, 단계 555로 진행한다.

단계 550에서 웹 서버는 단계 545에서 판단한 결과에 따라서 보장형 서비스에 대한 해당 서비스 정보를 제공한다.

예를 들면, 웹 서버는 사용자가 요구한 오디오 서비스 또는 비디오 서비스 등과 같은 정보와 같이 실시간으로 실행해야 하는 보장형 서비스에 대하여는 항상 우선적으로 사용자의 요구에 해당하는 서비스 정보를 제공한다.

단계 555에서 웹 서버는 현재의 보장형 서비스(예를 들면, 오디오 서비스 또는 비디오 서비스 등)의 수량을 알코라이즈를지를 판단하고, 보장형 서비스를 완료하였으면 단계 560으로 진행한다. 그러나 완료하지 않았다면, 다시 단계 550으로 진행한다.

단계 560에서 서비스 스케줄러부(315 - 도 3 참조)는 보장형 서비스가 완료 될때 다른 일정한 신호를 트래픽 부류별 토큰 발생부(310 - 도 3 참조)로 전달하고, 트래픽 부류별 토큰 발생부(310 - 도 3 참조)는 수신한 일정한 신호에 따라서 수량 판별할 해당 보장형 서비스의 토큰을 제거하고, 이 후 계속적으로 수신되는 사용자 요구에 해당하는 서비스를 제공하기 위해 단계 545로 진행한다.

예를 들어, 단계 545에서 단계 560까지를 설명하면, 사용자로부터 오디오 서비스의 요구가 수신된 경우, 웹 서버는 이러한 오디오 서비스에 대하여 미리 설정된 서비스 종류, 중, 저, 어디에 해당하는지를 판단한다. 여기서의 판단은 사용자 요구에 따른 패킷(Packet)의 헤더를 검색하여 이루어진다.

상기 검색 결과에 따라서, 웹 서버는 사용자로부터 수신된 서비스 요구가 오디오 서비스이고, 이러한 오디오 서비스는 보장형 서비스에 해당한다고 판단한다. 그리고 수신된 서비스 요구에 따라서 토큰(token)을 부여하고, 미리 설정된 상위 우선 순위의 서비스 제공 순서에 따라서 보장형 서비스에 대한 토큰을 저장하고 있는 토큰 배분 서비스 스케줄러부(315 - 도 3 참조)를 이용하여 검색하고, 존재하는 해당 서비스의 토큰에 따라 해당 서비스를 수행한다.

이러한 보장형 서비스의 수행 중에는 상위 우선 순위의 토큰 존재 여부를 판단할 필요가 없는 데, 이는 보장형 서비스가 가장 상위 우선 순위의 서비스이기 때문이다.

단계 565에서 서비스 스케줄러부(315 - 도 3 참조)는 부하 제어 서비스(예를 들면, 항공기 예약과 같은 온라인 서비스 등)에 대한 토큰 배분을 검사하여 부하 제어 서비스의 토큰이 존재하는지를 판단하고, 단계 565에서의 판단 결과, 부하 제어 서비스의 토큰이 존재한다면 단계 570으로 진행하고, 부하 제어 서비스의 토큰이 존재하지 않는다면 단계 530으로 진행한다.

단계 570에서 웹 서버는 사용자의 요구에 의한 해당 부하 제어 서비스(예를 들면, 항공기 예약과 같은 온라인 서비스 등) 제공을 수행한다.

단계 575에서 웹 서버는 토큰 발생부(310 - 도 3 참조)에 의해 현재 사용자 요구에 해당하는 상위 우선 순위의 토큰이 부여되었는지를 판단하고, 상위 우선 순위의 토큰이 부여되지 않았으면 단계 570으로 진행하고, 상위 우선 순위의 토큰이 부여되었으면 단계 545로 진행한다.

예를 들어, 단계 575를 설명하면, 웹 서버가 현재 항공기 예약과 같은 온라인 서비스(부하 제어 서비스에 해당함)를 수행하고 있는 도중에 사용자로부터 오디오 서비스 또는 비디오 서비스와 같은 보장형 서비스에 대한 요구가 수신된 경우, 이는 현재 수행 중인 부하 제어 서비스보다 상위 우선 순위의 서비스로 분류된 것으로서, 현재의 부하 제어 서비스의 수행을 중지하고, 보장형 서비스 요구에 따른 해당 서비스를 수행하게 된다.

그리고 단계 580에서 웹 서버는 사용자 요구에 의한 해당 부하 제어 서비스(예를 들면, 항공기 예약과 같은 온라인 서비스)의 완료 여부를 판단하여, 해당 부하 제어 서비스의 수행이 완료되었다면 단계 585로 진행하고, 그렇지 않다면 단계 570으로 진행한다.

단계 595에서 서비스 스케줄러부(315 - 도 3 참조)는 부하 제어 서비스가 완료됨에 따른 일정한 산출을 트래픽 부동할 토큰 발생부(310 - 도 3 참조)에 전달하고, 트래픽 부동할 토큰 발생부(310 - 도 3 참조)는 수신한 일정한 산출에 따라서 해당 부하 제어 서비스의 토큰을 제거한다. 이 후, 웹 서버는 단계 545로 진행하여 수신되는 사용자의 서비스 요구에 따라서 해당 서비스 정보를 계속하여 제공한다.

예를 들어 단계 595에서 단계 590까지를 설명하면, 사용자로부터 항공기 예약 등과 같은 서비스 요구가 수신된 경우, 웹 서버는 이러한 항공기 예약 서비스에 대하여 미리 설정된 서비스 종류 중 어디에 해당하는가를 판단한다. 여기서의 판단은 사용자 요구에 따른 패킷(Packet)의 헤더를 검색하여 이루어진다.

상기 검색 결과에 따라서, 웹 서버는 사용자로부터 수신된 서비스 요구가 항공기 예약 서비스이고, 이러한 항공기 예약 서비스는 부하 제어 서비스에 해당한다고 판단한다. 그리고 수신된 서비스 요구에 따라서 토큰(token)을 부여하고, 미리 설정된 상위 우선 순위의 서비스 제공 순서에 따라서 부하 제어 서비스에 대한 토큰을 제공하고 있는 토큰 버퍼를 서비스 스케줄러(315 - 도 3 참조)를 이용하여 검색하고, 존재하는 해당 서비스의 토큰에 따라 해당 서비스를 수행한다.

이러한 부하 제어 서비스의 수행 중에, 상위 우선 순위의 토큰 존재 여부를 판단하게 되는 데, 이는 부하 제어 서비스보다 상위 우선 순위의 보장형 서비스가 존재할 수 있기 때문이다.

따라서, 현재 부하 제어 서비스로서 항공기 예약 서비스를 제공하는 있는 도중에 오디오 서비스 또는 비디오 서비스 등과 같은 우선 순위가 높은 보장형 서비스에 대한 사용자 요구가 존재시에, 현재 수행 중인 항공기 예약 서비스보다 우선적으로 오디오 서비스 또는 비디오 서비스 등과 같은 보장형 서비스를 수행한다. 이렇게 보장형 서비스를 먼저 수행하고 나니, 현재 수행 중이던 부하 제어 서비스에 계속해서 수행한다.

단계 598에서 서비스 스케줄러부(215 - 도 2 참조)는 기존 인터넷 서비스(예를 들어, 파일 다운로드 또는 이메일 서비스 등)의 토큰 버퍼를 검사하여, 사용자의 요구에 따른 기존 인터넷 서비스에 해당하는 토큰이 존재하면, 단계 595로 진행한다.

그러나 기존 인터넷 서비스의 토큰이 존재하지 않는다면 단계 615로 진행한다.

단계 616에서 사용자로부터 수신된 서비스 요구는 미리 설정된 서비스 종류 중에서 어디에도 해당되지 않는다고 판단하고, 여러 메시지를 출력한 후에, 단계 630으로 진행하여 계속적으로 사용자로부터 서비스 요구를 수신한다.

단계 595에서 웹 서버는 사용자로부터의 기존 인터넷 서비스(예를 들어, 파일 다운로드 등) 요구에 해당 하는 서비스를 수행한다.

단계 600에서 웹 서버는 상위 우선 순위의 서비스 요구가 발생하였는지를 판단하는데, 이를 다시 말하면, 토큰 발생부(310 - 도 3 참조)에서 새로운 사용자 요구에 해당하는 상위 우선 순위의 토큰을 부여하였는지를 판단한다. 상위 우선 토큰을 부여하였다면 단계 545로 진행하고, 상위 우선 순위의 토큰이 부여되지 않았다면 단계 605로 진행한다.

예를 들어, 웹 서버는 파일 다운로드와 같은 기존 인터넷 서비스를 수행 중이라도 언제든지 상위 우선 순위의 토큰 가지고 있는 해당 서비스를 수행한다.

그리고 단계 605에서 웹 서버는 기존 인터넷 서비스의 완료 여부를 판단하여, 기존 인터넷 서비스가 완료 되었다면 단계 610으로 진행하고, 기존 인터넷 서비스가 완료되지 않았다면 기존 인터넷 서비스의 수행을 완료하기 위해 단계 595로 다시 진행한다.

단계 610에서 서비스 스케줄러부(315 - 도 3 참조)는 기존 인터넷 서비스 중에서 해당 요구 서비스에 할당됨에 따라 상기 요구 서비스에 해당하는 토큰을 트래픽 부동할 토큰 발생부(310 - 도 3 참조)로 전달하고, 트래픽 부동할 토큰 발생부(310 - 도 3 참조)는 제공이 완료된 기존 인터넷 서비스에 해당하는 토큰을 제거한다. 이 후, 웹 서버는 단계 545로 진행하여 수신되는 사용자의 서비스 요구에 해당하는 서비스 정보를 계속적으로 제공한다.

예를 들어 단계 600에서 단계 605까지를 설명하면, 사용자로부터 파일 다운로드, 이메일 등과 같은 서비스 요구가 수신된 경우, 웹 서버는 이러한 파일 다운로드, 이메일 등과 같은 서비스에 대하여 미리 설정된 서비스 종류 중 어디에 해당하는가를 판단한다. 여기서의 판단은 사용자 요구에 따른 패킷(Packet)의 헤더를 검색하여 이루어진다.

상기 검색 결과에 따라서, 웹 서버는 사용자로부터 수신된 서비스 요구가 이메일 서비스이고, 이러한 이메일 서비스는 기존 인터넷 서비스에 해당한다고 판단한다. 그리고 수신된 서비스 요구에 따라서 토큰(token)을 부여하고, 미리 설정된 상위 우선 순위의 서비스 제공 순서에 따라서 기존 인터넷 서비스에 대한 토큰을 제공하고 있는 토큰 버퍼를 서비스 스케줄러(315 - 도 3 참조)를 이용하여 검색하고, 존재하는 해당 서비스의 토큰에 따라 해당 서비스를 수행한다.

이러한 기존 인터넷 서비스의 수행 중에, 상위 우선 순위의 토큰 존재 여부를 판단하게 되는 데, 이는 기존 인터넷 서비스보다 상위 우선 순위의 보장형 서비스 또는 부하 제어 서비스가 존재할 수 있기 때문이다.

따라서, 현재 기존 인터넷 서비스로서 이메일 서비스를 제공하는 있는 도중에 오디오 서비스 또는 비디오 서비스 등과 같은 우선 순위가 높은 보장형 서비스에 대한 사용자 요구가 존재 시에도 우선 순위가 높은 부하 제어 서비스로서 항공기 예약과 같은 온라인 서비스에 대한 사용자 요구가 존재하는 경우도, 현재 수행 중인 이메일 서비스보다 오디오 서비스 또는 비디오 서비스 등과 같은 보장형 서비스를 우선적으로 수행한다. 이렇게 보장형 서비스를 먼저 수행하고 나니, 현재 수행 중이던 기존 인터넷 서비스를 계속해서 수행한다(새로운 요구 서비스가 부하 제어 서비스에 해당하는 경우에도 상기의 경우는 같다).

그리고 웹 서버는 상기와 같은 과정을 사용자에게 의해 서비스 요구가 계속해서 수신되는 한, 상기 과정을 반복적으로 계속해서 수행한다.

도 6a는 본 발명의 바람직한 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 서비스 요구 순서대로 데이터 처리를 간략하게 나타낸 도면이다.

도 6a를 참조하면, 단계 620에서 웹 서버는 도 2에서 나타난 바와 같이 웹 서버의 상태(도 1 참조)에 따라 제공 가능한 서비스를 설정한다.

다시 말하면, 정상 상태(100 - 도 1 참조) 및 최적 상태(105 - 도 1 참조)에서는 보정형 서비스, 부하 제어 서비스 및 기존 인터넷 서비스를 모두 제공하고, 가상의 혼잡 상태(110 - 도 1 참조)에서는 보정형 서비스 및 부하 제어 서비스를 제공하며, 그리고 혼잡 상태(115 - 도 1 참조)에서는 보정형 서비스만을 제공한다.

단계 625에서 웹 서버는 사용자의 서비스 요구를 수신하고, 단계 630으로 진행한다.

단계 630에서 웹 서버는 현재의 부하량을 실시간으로 측정하며, 도 1에서 설명된 정상 상태(100 - 도 1 참조), 최적 상태(105 - 도 1 참조), 가상의 혼잡 상태(110 - 도 1 참조) 그리고 혼잡 상태(115 - 도 1 참조) 인지를 검색한 후, 단계 635로 진행한다.

이 때, 상기의 네 가지 상태는 임의적으로 일정한 부하량에 따라 구분하여 적용할 수가 있다. 따라서, 설명의 편의상 부하량에 따른 각 상태를 예를 들면, 0퍼센트에서 30퍼센트 미만의 부하량이 존재하는 경우를 정상 상태, 및 최적 상태, 30퍼센트 이상에서 70퍼센트 미만의 부하량이 존재하는 경우를 가상의 혼잡 상태, 70퍼센트 이상의 부하량이 존재하는 경우를 혼잡 상태로 구분할 수 있다.

단계 635에서 웹 서버는 단계 630에서 특정한 웹 서버 상태(예를 들면, 가상의 혼잡 상태)로 검색되었다면, 사용자에게 의한 서비스 요구 서비스가 현재 웹 서버의 상태(예를 들면, 가상의 혼잡 상태)에서 제공 가능한 서비스인지를 판단한다.

단계 635에서의 판단 결과, 제공 가능한 서비스라면 단계 640으로 진행하고, 그렇지 않다면 상기 과정을 종료한다.

예를 들어, 단계 635를 설명하면, 사용자의 요구 서비스가 파일 다운로드 또는 이메일 등과 같은 기존 인터넷 서비스라면 가상의 혼잡 상태에서는 제공할 수 없는 서비스로 상기 서비스는 제공할 수 없게 된다. 그러나 예를 달리 하여, 사용자의 요구 서비스가 오디오 서비스 또는 비디오 서비스와 같은 보정형 서비스라면 현재 웹 서버 상태가 가상의 혼잡 상태라 하더라도 상기 서비스는 제공할 수가 있다고 판단하고, 이 후 단계 640으로 진행한다.

단계 640에서 각 웹 서버 상태에 따른 제공 가능한 서비스라면, 웹 서버는 사용자의 요구가 존재한 순서(서버 상태에 따라 다른 우선 순위를 적용하지 않음)대로 서비스를 수행한다.

도 6b는 본 발명의 바람직한 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 서비스 요구 순서대로 데이터 처리를 상세하게 나타낸 도면이다.

도 6b를 참조하면, 단계 645에서 웹 서버는 도 2에서 나타난 바와 같이 웹 서버의 상태(도 1 참조)에 따라 제공 가능한 서비스를 설정한다.

다시 말하면, 정상 상태(100 - 도 1 참조) 및 최적 상태(105 - 도 1 참조)에서는 보정형 서비스, 부하 제어 서비스 및 기존 인터넷 서비스를 모두 제공하고, 가상의 혼잡 상태(110 - 도 1 참조)에서는 보정형 서비스 및 부하 제어 서비스를 제공하며, 그리고 혼잡 상태(115 - 도 1 참조)에서는 보정형 서비스만을 제공한다.

여기서 보정형 서비스는 'IETF(Internet Engineering Task Force) RFC(request for comments) 2212'에 정의되어 있는 서비스로서, 이는 엄격한 QoS(Quality of Service)를 보장하는 서비스이다. 예를 들면, 정해진 데이터 레이트(Data rate)를 보장하고, 바운드드 딜레이(bounded delay), 노 큐잉 로스(no queueing loss) 등의 특징을 가지고 있어, 하드 리얼 타임 애플리케이션(hard real-time applications) 또는 서클 로 애플리케이션(circuit emulation) 등에 적용되는 것으로 오디오 서비스나 비디오 서비스 등이 있을 수 있다.

부하 제어용 서비스는 'RFC 2212'에 정의되어 있는 서비스이다. 예를 들면, 부하가 적은 때의 기존 인터넷 서비스나 근거리 서비스로서 대역폭, 지연 시간, 패킷 손실에 대한 품질 보장을 얻고, 네트워크 내의 총합한 자원을 확보함으로써, 높은 전송 확률을 제공한다. 따라서 이 서비스는 어댑티브 리얼 타임 애플리케이션(Adaptive real-time applications) 등에 적용되는 것으로 항공기 예약 서비스 등과 같은 온라인 서비스 등이 있을 수 있다.

그리고 기존 인터넷 서비스는 일반적인 사용자 서비스 요구로서 파일 다운로드, 이메일 서비스 등이 있을 수 있다.

단계 650에서 웹 서버는 사용자의 서비스 요구를 수신하고, 수신한 서비스 요구에 해당하는 트래픽을 서버 내의 트래픽 분류부(400 - 도 4 참조)로 전달하고, 트래픽 분류부(400 - 도 4 참조)는 사용자의 서비스 요구에 해당하는 트래픽을 분류한 후, 단계 655로 진행한다.

이러한 트래픽의 서버 내에서의 분류 방법은 패킷(Packet) 내 특정 필드의 해석으로 구현할 수 있다.

단계 655에서, 트래픽 분류부(400 - 도 4 참조)에 의해 분류된 트래픽(사용자에 의한 요구 서비스)에 대하여, 트래픽 분류부(400)로 발생된(405 - 도 4 참조)는 사용자의 서비스 요구에 해당하는 트래픽이 로드를 부여하고(도 4 단계 415를 증가시킴), 로드를 서비스 스케줄러부(415 - 도 4 참조)로 전달한다.

그리고 단계 660에서 부하 감시 장치부(410 - 도 4 참조)는 현재 웹 서버의 부하량을 실시간으로 검색한다. 이는 현재 웹 서버 상태가 단계 645에서 설정된 네 가지의 상태 중에서 어디에 해당하는지를 체크하

는 것이다.

단계 65에서 웹 서버는 단계 650에서 부하 감시 장치부(410 - 도 4 참조)에 의해 검색된 결과에 따라서 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)인지를 판단한다.

단계 66에서 웹 서버의 판단 결과, 웹 서버의 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)라면, 단계 670으로 진행한다. 그러나 그렇지 않다면 단계 680으로 진행된다.

여기서 단계 665에서의 판단 결과, 웹 서버의 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)라면, 제공 가능한 서비스는 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스를 의미한다. 그리고 이렇게 제공되는 서비스는 사용자의 요구 순서대로 이루어진다.

단계 670에서 웹 서버는 서비스 스케줄러부(415 - 도 4 참조)에 의해 관리되는 토콘 배치를 검사하여 사용자의 서비스 요구 순서대로 부여된 토콘에 따라 해당 서비스를 제공한다.

예를 들면, 사용자의 서비스 요구가 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 순서에 상관없이 어느 서비스 요구이든 시간 우선적으로 먼저 요구된 서비스를 수행한다.

그리고 단계 675에서 웹 서버는 수행 중인 서비스의 완료 여부를 판단한다. 단계 675에서의 판단 결과, 서비스 수행이 완료되었다면, 상기 과정은 종료하고, 서비스 수행이 완료되지 않았다면 단계 680으로 진행한다.

단계 680에서 웹 서버는 사용자에게 의해 요구된 서비스 주행이 일정 시간(예를 들면, 파일 다운로드와 같은 경우는 1시간 등)을 해당 서비스에 따라 일정한 시간을 정하고, 이렇게 정하는 이유는 지정된 시간 내에 서비스가 수행되지 않을에도 불구하고 계속 수행한다면 다른 사용자의 서비스 요구에는 서비스 제공이 불가능하기 때문이다)이 초과되었는지를 판단하고, 지정된 시간(예를 들면, 1시간)이 초과된 경우에는 상기 과정을 종료하고, 지정된 시간(예를 들면, 1시간)이 초과되지 않았다면 단계 670으로 진행된다.

단계 685에서 웹 서버는 단계 660에서 부하 감시 장치부(410 - 도 4 참조)에 의해 검색된 결과에 따라서 현재 상태가 가상의 혼잡 상태(110 - 도 1 참조)인지를 판단한다.

단계 685에서의 판단 결과, 웹 서버의 현재 상태가 가상의 혼잡 상태(110 - 도 1 참조)라면, 단계 690으로 진행된다. 그러나 그렇지 않다면 단계 700으로 진행된다.

여기서 웹 서버의 현재 상태가 가상의 혼잡 상태(110 - 도 1 참조)라면 제공 가능한 서비스는 보장형 서비스와 부하 제어 서비스이다. 그리고 이들 서비스는 사용자의 요구 순서대로, 사용자에게 제공한다.

단계 690에서 웹 서버는 사용자의 요구 서비스가 기존 인터넷 서비스인지를 판단한다. 단계 690에서의 판단 결과, 기존 인터넷 서비스라면 단계 695로 진행하고, 기존 인터넷 서비스가 아니라면 단계 700으로 진행한다.

단계 695에서 웹 서버는 단계 645에서 설정된 조건에 따라 가상의 혼잡 상태(110 - 도 1 참조)에서 기존 인터넷 서비스를 제공할 수 없기에 서비스 불가 메시지에 사용자에게 전달하고 상기 과정을 종료한다.

단계 700에서 웹 서버는 현재의 웹 서버 상태가 혼잡 상태(115 - 도 1 참조)인지를 판단한다. 단계 700에서의 판단 결과 혼잡 상태라면 단계 705로 진행하고, 그렇지 않다면 단계 660으로 진행된다.

여기서 웹 서버는 혼잡 상태(115 - 도 1 참조)에서 제공할 수 있는 서비스는 보장형 서비스이다.

단계 705에서 웹 서버는 사용자의 요구 서비스가 보장형 서비스인지를 판단하여, 그렇다면 단계 670으로 진행하고, 그렇지 않다면 단계 695로 진행된다.

이렇게 보 말하면 따르면, 웹 서버는 특정한 웹 서버 상태에 따라 제공 가능한 서비스를 미리 설정하고, 사용자의 요구 서비스 발생시에 사용자의 서비스 요구 순서대로 서비스를 제공한다.

그리고 상기 과정은 사용자로부터 서비스 요구가 수신되는 한 계속적으로 반복하여 수행된다.

도 7a는 본 발명의 바람직한 또 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 요구 서비스에 우선 순위를 적용한 데이터 처리를 간략하게 나타낸 도면이다.

도 7a를 참조하면, 단계 710에서 웹 서버는 도 2에서 나타난 바와 같이 웹 서버의 상태(도 1 참조)에 따라 제공 가능한 서비스 및 제공 가능한 서비스 간에 사용자에게 제공되는 우선 순위를 설정한다.

다시 말하면, 정상 상태(100 - 도 1 참조) 및 최적 상태(105 - 도 1 참조)에서는 보장형 서비스, 부하 제어 서비스 및 기존 인터넷 서비스를 모두 제공하고, 가상의 혼잡 상태(110 - 도 1 참조)에서는 보장형 서비스 및 부하 제어 서비스를 제공하며, 그리고 혼잡 상태(115 - 도 1 참조)에서는 보장형 서비스를 제공한다.

또한 웹 서버는 각 제공 가능한 서비스 간에는 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 순서로 우선 순위를 분류하고, 사용자의 요구 순서에 상관없이 상기 설정된 우선 순위에 따라 사용자에게 서비스를 제공한다.

단계 715에서 웹 서버는 사용자의 서비스 요구를 수신하고 단계 720으로 진행한다.

그리고 단계 720에서 웹 서버는 현재의 웹 서버 상태를 검색하고, 단계 725로 진행한다.

단계 725에서 웹 서버는 사용자 요구 서비스가 현재의 웹 서버 상태에 따른 제공 가능한 서비스인지를 판단한다.

예를 들면, 현재의 웹 서버 상태가 가상의 혼잡 상태(110 - 도 1 참조)라면, 제공 가능한 서비스는 보장형 서비스 그리고 부하 제어 서비스이고, 이들 서비스간에는 단계 710에서 설정된 우선 순위에 따라 보장형 서비스를 우선 제공하고 나서, 부하 제어 서비스를 제공한다. 따라서, 웹 서버는 사용자 요구 서비스

가 상기에 언급한 서비스에 해당하는지를 판단하는 것이다.

단계 73에서 웹 서버는 단계 72에서의 판단 결과에 따라서, 사용자에게 의한 요구 서비스를 제공한다.

상기 도 7a를 참조한 설명은 본 발명의 내용을 간략히 설명한 것으로서, 도 7b와 도 7c를 통하여 더욱 상세하게 설명하기로 한다.

도 7b와 도 7c는 본 발명의 바람직한 또 다른 실시예에 따른 웹 서버 상태를 바탕으로 사용자의 요구 서비스에 우선 순위를 적용한 데이터 처리를 상세하게 나타낸 도면.

하기에서는 설명의 편의상 도 7b와 도 7c를 함께 참조하여 설명하기로 한다.

도 7b와 도 7c를 참조하면, 단계 73에서 웹 서버는 도 2에서 나타낸 바와 같이 웹 서버의 상태(도 1 참조)에 따라 제공 가능한 서비스 및 제공 가능한 서비스 간에 사용자에게 제공되는 우선 순위를 설정한다.

다시 말하면, 정상 상태(100 - 도 1 참조) 및 최적 상태(105 - 도 1 참조)에서는 보정형 서비스, 부하 제어 서비스 및 기타 여러 서비스를 모두 제공하고, 가상의 혼합 상태(110 - 도 1 참조)에서는 보정형 서비스 및 부하 제어 서비스를 제공하고, 그리고 혼합 상태(115 - 도 1 참조)에서는 보정형 서비스만을 제공한다.

이 때, 상기의 네 가지 상태는 임의적으로 일정한 부하량에 따라 구분하여 적용할 수가 있다. 따라서, 설명의 편의상 부하량에 따른 각 상태를 예를 들면, 0퍼센트에서 30퍼센트 미만의 부하량에 존재하는 경우를 정상 상태 및 최적 상태, 30퍼센트 이상에서 70퍼센트 미만의 부하량에 존재하는 경우를 가상의 혼합 상태, 70퍼센트 이상의 부하량에 존재하는 경우를 혼합 상태로 구분할 수 있다.

또한 웹 서버는 각 제공 가능한 서비스 간에는 보정형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 순서로 우선 순위를 분류하고, 사용자의 요구 순서에 상관없이 상기 설정된 우선 순위에 따라 사용자에게 서비스를 제공한다.

단계 740에서 웹 서버는 사용자의 서비스 요구를 수신하고, 수신한 서비스 요구에 해당하는 트래픽을 웹 서버 내의 트래픽 분류부(400 - 도 4 참조)로 전달하고, 트래픽 분류부(400 - 도 4 참조)는 사용자의 서비스 요구에 해당하는 트래픽을 분류한 후, 단계 745로 진행한다.

이러한 트래픽(사용자에 의한 요구 서비스)의 서버 내에서의 분류 방법은 패킷(Packet) 내 특정 필드의 형식으로 구현할 수 있다.

단계 745에서 트래픽 분류부(400 - 도 4 참조)에 의해 분류된 트래픽(사용자에 의한 요구 서비스)에 대하여 트래픽 분류부(400 - 도 4 참조)는 사용자의 요구 서비스에 해당하는 트래픽에 토른을 부여하여(토른 수 1을 증가시킴), 토른을 서비스 스케줄러부(415 - 도 4 참조)로 전달한다.

그리고 단계 750에서 부하 감시 장치부(410 - 도 4 참조)는 현재 웹 서버의 부하량을 실시간으로 검색한 뒤, 이는 현재 웹 서버 상태가 단계 73에서 설정된 네 가지의 상태 중에서 어디에 해당하는지를 체크하는 것이다.

단계 755에서 웹 서버는 단계 750에서 부하 감시 장치부(410 - 도 4 참조)에 의해 검색된 결과에 따라서 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)인지를 판단한다.

단계 755에서의 판단 결과, 웹 서버의 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)라면, 단계 760으로 진행한다. 그러나 그렇지 않다면 단계 765로 진행한다.

여기서 단계 755에서의 판단 결과, 웹 서버의 현재 상태가 정상 상태(100 - 도 1 참조) 또는 최적 상태(105 - 도 1 참조)라면, 제공 가능한 서비스는 보정형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스를이다. 그리고 이렇게 제공되는 서비스는 보정형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스의 우선 순위가 적용되어 사용자에게 제공된다.

여기서 보정형 서비스는 "IETF(Internet Engineering Task Force) RFC(request for comments) 2212에 정의되어 있는 서비스로서 이는 엄격한 QoS(Quality of Service)를 보장하는 서비스이다. 예를 들면, 정해진 데이터 레이트(Data rate)를 보장하고, 바운디드 딜레이(bounded delay), 노 큐잉 로스(no queuing loss) 등의 특성을 가지고 있어, 하드 리얼 타임 애플리케이션(hard real-time applications) 또는 서킷 에뮬레이션(circuit emulation) 등에 적용되는 것으로 오디오 서비스나 비디오 서비스 등이 있을 수 있다.

그리고 부하 제어형 서비스는 "RFC 2212"에 정의되어 있는 서비스이다. 예를 들면, 부하가 적을 때의 기존 인터넷 서비스와 근사한 서비스로서 대역폭, 지연 시간, 패킷 손실에 대한 절대적 보장이 없으나, 네트워크 내의 충분한 자원을 확보함으로써, 높은 전송 확률을 제공한다. 따라서 이 서비스는 소프트 리얼 타임 애플리케이션(Adaptive real-time applications) 등에 적용되는 것으로 항공기 예약 서비스 등과 같은 온라인 서비스 등이 있을 수 있다.

그리고 기존 인터넷 서비스는 일반적인 사용자 서비스 요구로서 파일 다운로드, 이메일 서비스 등이 있을 수 있다.

단계 760에서 서비스 스케줄러(415 - 도 4 참조)는 사용자의 요구 서비스에 해당하는 토른에 조정되어 있는 토른 배분(보정형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스)에 해당하는 토른 배분)을 검색하고 단계 765로 진행한다.

단계 765에서 웹 서버는 단계 760에서 검색한 토른에 해당하는 서비스가 보정형 서비스인지지를 판단한다. 상기에서 검색한 토른에 해당하는 서비스가 보정형 서비스라면, 단계 770으로 진행하고, 그렇지 않다면 단계 785로 진행한다.

단계 770에서 웹 서버는 사용자 요구에 의한 서비스 즉, 오디오 서비스 또는 비디오 서비스 등을 수행한다.

단계 775에서 웹 서버는 단계 770에서 수행한 보장할 서비스가 완료되었는지를 판단하고, 수행이 완료되었다면 단계 780으로 진행하고, 수행이 완료되지 않았다면 단계 770으로 진행한다.

그리고 단계 780에서 사용자 요구에 의한 해당 보장할 서비스(예를 들면, 오디오 서비스 또는 비디오 서비스 등)의 수행 완료에 따라, 해당 서비스에 부여된 토큰을 제거하고 단계 740으로 진행한다.

단계 785에서 웹 서버는 현재 상태가 가장의 혼잡 상태(110 - 도 1 참조)인지를 판단한다. 여기서 가장의 혼잡 상태(110 - 도 1 참조)는 80퍼센트 이상에서 70퍼센트 미만인 부하량이 존재하는 경우로서, 제공 가능한 서비스는 보장할 서비스와 부하 제어 서비스이다.

단계 790에서 웹 서버 내의 서비스 스케줄러(415 - 도 4 참조)는 사용자의 요구 서비스에 해당하는 토큰이 저장되어 있는 토큰 버퍼, 즉 보장할 서비스에 해당하는 토큰 버퍼와 부하 제어 서비스에 해당하는 토큰 버퍼를 검색하고, 단계 765로 진행한다.

단계 795에서 웹 서버는 사용자의 요구에 해당하는 서비스가 부하 제어 서비스인지를 판단하여, 부하 제어 서비스라면 단계 800으로 진행하고, 그렇지 않다면 단계 830으로 진행한다.

단계 800에서 웹 서버는 단계 795의 판단 결과에 따른 부하 제어 서비스와 현재의 웹 서버 상태, 즉 가장의 혼잡 상태에서 제공 가능한 서비스인지를 판단하여, 제공 가능한 서비스라면 단계 805로 진행하고, 그렇지 않다면 단계 855로 진행한다.

단계 805에서 웹 서버는 현재의 서버 상태에서 제공 가능한 할공기 여역 등과 같은 부하 제어 서비스를 수행한다.

단계 810에서 웹 서버는 현재 수행 중인 부하 제어 서비스가 완료되었는지를 판단하여 완료되었다면 단계 815로 진행하고, 그렇지 않다면 단계 805로 진행한다.

그리고 단계 815에서 토큰 발생부(405 - 도 4 참조)는 사용자 요구에 의한 해당 부하 제어 서비스(예를 들면, 할공기 여역 등과 같은 서비스)의 수행 완료에 따라, 해당 서비스에 부여된 토큰을 제거하고 단계 740으로 진행한다.

단계 820에서 웹 서버는 현재 상태가 혼잡 상태(115 - 도 1 참조)인지를 판단한다. 여기서 혼잡 상태(115 - 도 1 참조)는 70퍼센트 이상의 부하량이 존재하는 경우로서, 제공 가능한 서비스는 보장할 서비스이다.

웹 서버는 단계 820에서의 판단에 따라서, 현재의 상태가 혼잡 상태(115 - 도 1 참조)라면 단계 825로 진행하고, 그렇지 않다면 단계 750으로 진행한다.

단계 825에서 웹 서버는 사용자 요구에 의한 보장할 서비스에 해당하는 토큰 버퍼를 검색하고, 단계 765로 진행한다.

그리고 단계 830에서 웹 서버는 사용자 요구에 해당하는 서비스가 파일 다운로드 등과 같은 기존 인터넷 서비스인지를 판단하여, 기존 인터넷 서비스라면 단계 835로 진행하고, 그렇지 않다면 단계 865로 진행한다.

단계 835에서 웹 서버는 사용자 요구에 해당하는 서비스가 파일 다운로드 등과 같은 기존 인터넷 서비스가 현재의 서버 상태에서 제공 가능한 서비스인지를 판단하여, 제공 가능한 서비스라면 단계 840으로 진행하고, 그렇지 않다면 단계 865로 진행한다.

단계 840에서 웹 서버는 사용자 요구에 의한 서비스 즉, 파일 다운로드 등과 같은 기존 인터넷 서비스를 수행한다.

단계 845에서 웹 서버는 단계 840에서 수행한 기존 인터넷 서비스가 완료되었는지를 판단하고, 수행이 완료되었다면 단계 850으로 진행하고, 수행이 완료되지 않았다면 단계 840으로 진행한다.

그리고 단계 850에서 사용자 요구에 의한 해당 기존 인터넷 서비스(예를 들면, 파일 다운로드 등)의 수행 완료에 따라, 해당 서비스에 부여된 토큰을 제거하고 단계 740으로 진행한다.

단계 855에서 웹 서버는 제공이 불가능한 서비스인 경우에 해당하는 여러 메시지를 출력하고, 단계 740으로 진행한다.

사용자로부터 서비스 요구가 수신되는 한, 웹 서버는 상기 과정을 계속적으로 반복하여 실행한다.

본 발명은 상기 실시예에 한정되지 않으며, 많은 변형이 본 발명의 사상 내에서 통상의 지식을 가진 자에 의하여 가능함은 물론이다.

발명의 효과

본 발명에 따라 인터넷 상에서 웹 서비스를 제공하고 있는 웹 서버의 운영을 실시간 트래픽 감시 정보를 기초로 하여 트래픽 관리 기능을 적용함으로써 기존의 하드웨어적인 업그레이드로만 증가하는 웹 서버에의 사용자 요구를 지원하던 운영 방법을 소프트웨어적 혹은 알고리즘적으로 해결할 수 있다.

그리고 하드웨어 업그레이드라는 방법이 극복할 수 있는 수준을 한 층 더 뛰어 넘어 기존 비동기 전달 방식(ATM; Asynchronous Transfer Mode) 및 패킷 교환 망에서 사용한 트래픽 제어 및 관리 방법을,

웹 서버 운용에 적합함으로써 지속적으로 증가하는 웹 서비스에 대한 차별화된 서비스를 요구하는 사용자 요구 사항에 부응하여 사용자에게 합리적인 가격의 서비스를 제공할 수 있다.

(57) 청구의 범위

청구항 1

웹 서비스를 위한 웹 서버 운용 방법에 있어서,
 사용자로부터 웹 서비스 요청 정보를 수신하는 단계;
 상기 수신된 웹 서비스의 종류를 판별하는 단계;
 상기 판별 결과, 미리 설정된 서비스의 우선 순위에 따라 해당 서비스를 수행하는 단계를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 2

제1항에 있어서,
 상기 웹 서비스는 보장형 서비스, 부하 제어형 서비스 그리고 인터넷 서비스로 구분하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 3

제1항에 있어서,
 상기 서비스는 요청된 각 서비스마다 토큰을 부여하고, 토큰이 부여된 각 서비스에 대해 상기 설정된 우선 순위에 따라 수행되는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 4

웹 서비스를 위한 웹 서버 운용 방법에 있어서,
 웹 서비스의 종류에 따라 우선 순위를 설정하는 단계;
 사용자로부터 웹 서비스 요청이 있는 경우 해당 웹 서비스를 수행하는 단계;
 상기 웹 서비스 수행 중에 다른 종류의 웹 서비스 정보 요청이 있는 경우, 요청된 웹 서비스의 종류를 판별하여 진행 중인 웹 서비스와 현재 요청된 웹 서비스의 우선 순위를 비교 판단하는 단계;
 상기 판단 결과, 우선 순위가 높은 웹 서비스를 수행하는 단계를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 5

제4항에 있어서,
 상기 웹 서비스는 보장형 서비스, 부하 제어형 서비스 그리고 인터넷 서비스로 구분하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 6

제4항에 있어서,
 상기 서비스는 요청된 각 서비스마다 토큰을 부여하고, 토큰이 부여된 각 서비스에 대해 상기 설정된 우선 순위에 따라 수행되는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 7

특수의 서비스에 대해 각각 우선 순위를 설정하는 단계;
 사용자로부터 서비스 요청 정보를 수신하는 단계;

상기 수산한 서비스 요청 정보에 상응하여 토큰을 부여하는 단계;

상기 부여한 토큰에 해당하는 서비스를 제공하고, 상기 서비스의 제공 중에 상위 우선 순위의 토큰이 존재하는 서비스의 요청이 있는지를 판단하는 단계;

상기 판단 결과, 제공 중인 서비스의 토큰보다 현재 요청된 서비스의 토큰이 우선 순위가 높은 경우, 우선 순위가 높은 해당 서비스를 우선적으로 사용자에게 제공하고, 이후에 제공 중인 서비스를 계속 수행하는 단계;

상기 서비스를 제공한 후, 해당 서비스 정보의 토큰을 제거하는 단계

를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 8

제7항에 있어서,

상기 판단 결과, 제공 중인 서비스의 토큰이 현재 요청된 서비스의 토큰보다 상기 우선 순위의 토큰인 경우에는 제공 중인 서비스를 계속 수행하는 단계를 더 포함하는 것

을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 9

제7항에 있어서,

상기 서비스는 보장형 서비스, 부하 제어 서비스 그리고 기존 인터넷 서비스 인 것

을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 10

특수의 서비스에 대하여 웹 서버 상태에 따른 우선 순위를 설정하는 단계;

사용자로부터 서비스 요청 정보를 수신하는 단계;

상기 수신한 서비스 요청 정보를 바탕으로 현재 웹 서버 상태를 검색하는 단계;

상기 검색한 웹 서버 상태에 따라 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인지를 판단하는 단계;

상기 판단 결과, 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인 경우에는 설정된 우선 순위에 따라 해당 서비스를 제공하는 단계

를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 11

제10항에 있어서,

상기 웹 서버 상태는 정상 상태, 최적 상태, 가상의 혼잡 상태, 혼잡 상태 중 적어도 어느 하나 인 것

을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 12

제11항에 있어서,

상기 웹 서버 상태는 부하량을 이용하여 설정하는 것

을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 13

제10항에 있어서,

상기 특수의 서비스는 보장형 서비스, 부하 제어 서비스, 기존 인터넷 서비스 중 적어도 어느 하나 인 것

을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 14

제10항에 있어서,

상기 복수의 서비스의 우선 순위는 보장형 서비스, 부하 제어 서비스, 기존 인터넷 서비스의 순서로 이루어진 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 15

제11항에 있어서,

상기 정상 상태 또는 최적 상태에서 제공 가능한 서비스는 보장형 서비스, 부하 제어 서비스, 기존 인터넷 서비스 중 적어도 어느 하나인 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 16

제11항에 있어서,

상기 가상의 혼잡 상태에서 제공 가능한 서비스는 보장형 서비스, 부하 제어 서비스 중 적어도 어느 하나인 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 17

제11항에 있어서,

상기 혼잡 상태에서 제공 가능한 서비스는 보장형 서비스인 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 18

제10항에 있어서,

상기 수신한 서비스 요청 정보에 해당하는 서비스에 토큰을 부여하는 단계를 더 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 19

제10항에 있어서,

상기 사용자에게 해당 서비스를 제공한 후, 제공 서비스에 해당하는 토큰을 제거하는 단계를 더 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 방법.

청구항 20

프로그램이 저장되어 있는 메모리;

상기 메모리에 저장되어 상기 프로그램을 실행하는 프로세서를 포함하되,

상기 프로세서는 상기 프로그램에 상응하여,

웹 서비스를 위한 웹 서버를 운용하는 데 있어서,

사용자로부터 웹 서비스 요청 정보를 수신하는 단계;

상기 수신된 웹 서비스의 종류를 판별하는 단계;

상기 판별 결과, 미리 설정된 서비스의 우선 순위에 따라 해당 서비스를 수행하는 단계를 실행하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 장치.

청구항 21

웹 서비스를 위한 웹 서버 운용 시스템에 있어서,
 사용자로부터 웹 서비스 요청 정보를 수신하는 수신부;
 상기 수신된 웹 서비스의 종류를 판별하는 판단부;
 상기 판별 결과, 미리 설정된 서비스의 우선 순위에 따라 해당 서비스를 수행하는 수행부
 를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 시스템.

청구항 22

프로그램이 저장되어 있는 메모리;
 상기 메모리에 저장되어 상기 프로그램을 실행하는 프로세서
 를 포함하되,
 상기 프로세서는 상기 프로그램에 상응하여,
 웹 서비스를 위한 웹 서버를 운용하는 데 있어서,
 웹 서비스의 종류에 따라 우선 순위를 설정하는 단계;
 사용자로부터 웹 서비스 요청이 있는 경우 해당 웹 서비스를 수행하는 단계;
 상기 웹 서비스 수행 중에 다른 종류의 웹 서비스 정보 요청이 있는 경우, 요청된 웹 서비스의 종류를 판
 별하여 진행 중인 웹 서비스와 현재 요청된 웹 서비스의 우선 순위를 비교 판단하는 단계;
 상기 판단 결과, 우선 순위가 높은 웹 서비스를 수행하는 단계
 를 실행하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 장치.

청구항 23

웹 서비스를 위한 웹 서버 운용 시스템에 있어서,
 웹 서비스의 종류에 따라 우선 순위를 설정하는 설정부;
 사용자로부터 웹 서비스 요청이 있는 경우 해당 웹 서비스를 수행하는 수행부;
 상기 웹 서비스 수행 중에 다른 종류의 웹 서비스 정보 요청이 있는 경우, 요청된 웹 서비스의 종류를 판
 별하여 진행 중인 웹 서비스와 현재 요청된 웹 서비스의 우선 순위를 더 비교 판단하는 판단부;
 상기 수행부는 상기 비교 판단의 결과에 따라 우선 순위가 높은 웹 서비스를 더 수행하는 것
 을 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 시스템.

청구항 24

프로그램이 저장되어 있는 메모리;
 상기 메모리에 저장되어 상기 프로그램을 실행하는 프로세서
 를 포함하되,
 상기 프로세서는 상기 프로그램에 상응하여,
 복수의 서비스에 대해 우선 순위를 설정하는 단계;
 사용자로부터 서비스 요청 정보를 수신하는 단계;
 상기 수신한 서비스 요청 정보에 상응하여 토권을 부여하는 단계;
 상기 부여한 토권에 해당하는 서비스를 제공하고, 상기 서비스의 제공 중에 상위 우선 순위의 토권이 존
 재하는 서비스의 요청이 있는지를 판단하는 단계;
 상기 판단 결과, 제공 중인 서비스의 토권보다 현재 요청된 서비스의 토권이 우선 순위가 높은 경우, 우
 선 순위가 높은 해당 서비스를 우선적으로 사용자에게 제공하고, 이후에 제공 중인 서비스를 계속 수행하
 는 단계;
 상기 서비스를 제공한 후, 해당 서비스 정보의 토권을 재가하는 단계
 를 실행하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 장치.

청구항 25

복수의 서비스에 대해 우선 순위를 설정하는 설정부;

사용자로부터 서비스 요청 정보를 수신하는 수신부;

상기 수신한 서비스 요청 정보에 상응하여 토큰을 부여하는 토큰 발행부;

상기 부여한 토큰에 해당하는 서비스를 제공하고, 상기 서비스의 제공 중에 상위 우선 순위의 토큰이 존 재하는 서비스의 요청이 있는지를 판단하는 서비스 스케줄러부;

상기 판단 결과, 제공 중인 서비스의 토큰보다 현재 요청된 서비스의 토큰이 우선 순위가 높은 경우, 우 선 순위가 높은 해당 서비스를 우선적으로 사용자에게 제공하고, 이후에 제공 중인 서비스를 계속 수행하 는 수행부

를 구비하되,

상기 토큰 발행부는 해당 서비스 정보의 토큰을 제거하는 것

을 더 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 시스템.

형구항 26

프로그램이 저장되어 있는 메모리;

상기 메모리에 저장되어 상기 프로그램을 실행하는 프로세서

를 포함하되,

상기 프로세서는 상기 프로그램에 상응하여,

복수의 서비스에 대하여 웹 서버 상태에 따른 우선 순위를 설정하는 단계;

사용자로부터 서비스 요청 정보를 수신하는 단계;

상기 수신한 서비스 요청 정보를 바탕으로 현재 웹 서버 상태를 검색하는 단계;

상기 검색한 웹 서버 상태에 따라 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인지를 판 단하는 단계;

상기 판단 결과, 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인 경우에는 설정된 우선 순위에 따라 해당 서비스를 제공하는 단계

를 실행하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 장치.

형구항 27

복수의 서비스에 대하여 웹 서버 상태에 따른 우선 순위를 설정하는 설정부;

사용자로부터 서비스 요청 정보를 수신하는 수신부;

상기 수신한 서비스 요청 정보를 바탕으로 현재 웹 서버 상태를 검색하는 검색부;

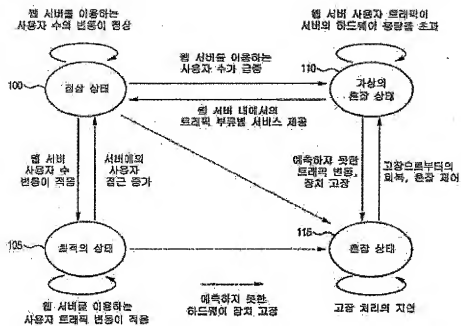
상기 검색한 웹 서버 상태에 따라 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인지를 판 단하는 판단부;

상기 판단 결과, 상기 수신한 서비스 요청 정보가 제공 가능한 서비스인 경우에는 설정된 우선 순위에 따라 해당 서비스를 제공하는 수행부

를 포함하는 것을 특징으로 하는 통합 웹 서비스를 위한 웹 서버 운용 시스템.

도면

도면1

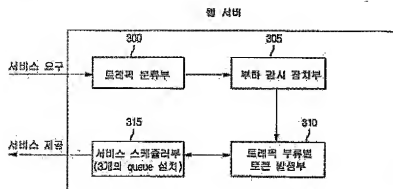


도면2

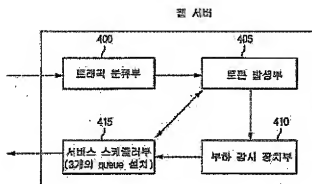
200

상 태	제공되는 서비스
최적 상태	A, B, C
정상 상태	A, B, C
가상의 손상 상태	A, B
손상 상태	A

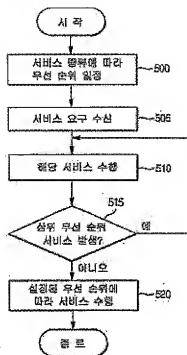
도면3



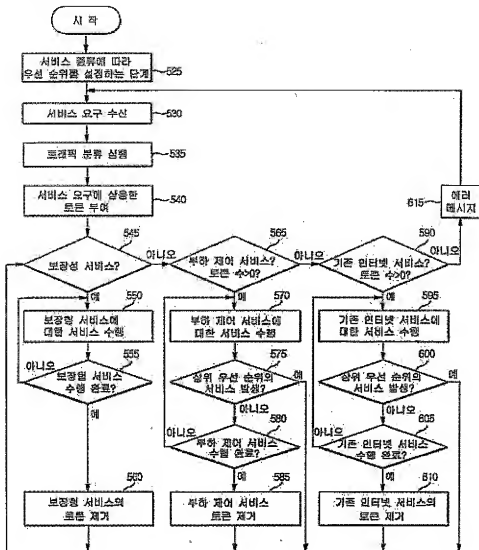
도면4



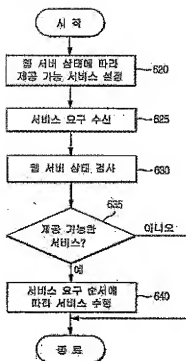
도면5a



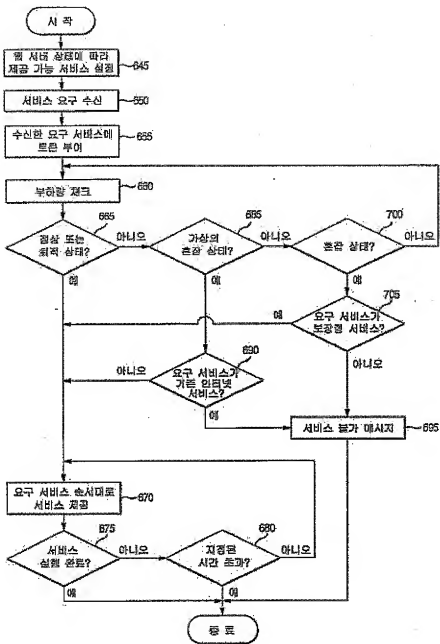
도 95b



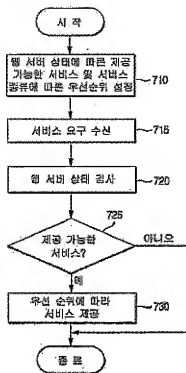
도 10a

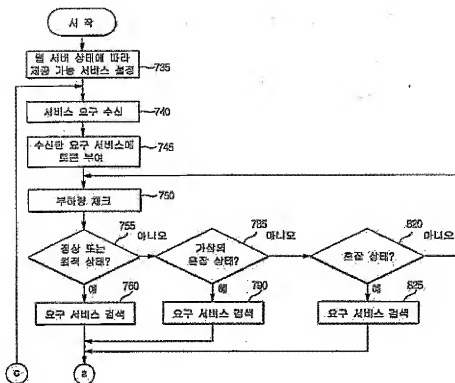


도 2000b



도 7c





도면 7e

